PLOS ONE

Characterisation of Major Histocompatibility Complex Class I in the Australian Cane Toad, *Rhinella marina*

Mette Lillie[1], Richard Shine[2], Katherine Belov[1]*

1 Faculty of Veterinary Science, University of Sydney, Sydney, New South Wales, Australia, 2 School of Biological Sciences, University of Sydney, Sydney, New South Wales, Australia

Abstract

The Major Histocompatibility Complex (MHC) class I is a highly variable gene family that encodes cell-surface receptors vital for recognition of intracellular pathogens and initiation of immune responses. The MHC class I has yet to be characterised in bufonid toads (Order: Anura; Suborder: Neobatrachia; Family: Bufonidae), a large and diverse family of anurans. Here we describe the characterisation of a classical MHC class I gene in the Australian cane toad, *Rhinella marina*. From 25 individuals sampled from the Australian population, we found only 3 alleles at this classical class I locus. We also found large number of class I alpha 1 alleles, implying an expansion of class I loci in this species. The low classical class I genetic diversity is likely the result of repeated bottleneck events, which arose as a result of the cane toad's complex history of introductions as a biocontrol agent and its subsequent invasion across Australia.

Introduction

The cane toad (*Rhinella marina*, formerly *Bufo marinus*) is a large, toxic anuran native to Mexico, Central and tropical South America, and the island of Trinidad [1]. The family Bufonidae is large and widespread, including 485 species of 35 genera, with representative species naturally occurring on all continents, except Australasia and Antarctica [2]. During the 1930s, the cane toad was a popular biological control agent for invertebrate agricultural pests and was introduced to numerous countries including Australia [1]. One hundred and one individuals were introduced in 1935 and bred for distribution to agricultural districts of north-eastern Australia [1]. They quickly became well established and began invading into northern and eastern Australia [1]. The cane toad population in Australia currently covers more than 1 million square kilometres and continues to spread [3].

The Major Histocompatibility Complex (MHC) is a large gene complex present in all jawed vertebrates with an integral role in the immune system. The antigen-presenting molecules encoded by the MHC class I and class II genes are cell-surface glycoproteins that bind intracellular and extracellular peptides, respectively [4]. The class I molecule consists of an α chain with 3 extracellular domains and a β2 microglobulin. Peptides derived from the degradation of intracellular pathogens are bound into a peptide-binding groove formed by the α1 and α2 domains [5]. Anchor residues in the peptide binding groove determine the specific peptides that can be bound by a particular MHC class I molecule. These peptide-MHC complexes are presented to cytotoxic T-lymphocytes and an appropriate immune response is launched on recognition of foreign bound peptides [4].

The MHC genes evolve under a birth and death model [6]. Multiple gene copies are generated through duplication events ("birth"). These two copies are then able to diverge and assume distinct peptide binding specificities or entirely new functions. Some of the duplicates may become pseudogenized by nonsense or frameshift mutations and "die". Through the course of this gene evolution, the numbers of MHC loci vary greatly within different species.

The class I genes are categorised as either classical or non-classical based on their functional differences. Classical MHC class I molecules are expressed on the surface of all nucleated cells, are highly polymorphic and are involved in presentation of self and non-self peptides to cytotoxic T cells [4]. The non-classical class I loci have conventionally been described as having variable expression in tissues, typically have low polymorphism, and have a variety of functions within the immune system [7]; although studies of classical and non-classical class I loci in human and mouse are challenging these definitions [8].

In anurans, it has long been held that there is clear distinction between the classical and non-classical loci. In the model anurans, *Xenopus laevis* and *Silurana tropicalis*, a single classical class I locus has been characterised in each species, along with a large expansion of non-classical class I loci [9,10,11,12]. These non-classical genes have low polymorphism and differ in their invariant amino acids in the peptide-binding groove, which determines their peptide anchoring [13]. A single classical class I locus was also identified in *Rana pipiens* [14] and *R. temporaria* [15]. But recently, multiple putatively classical class I loci have been identified in several anuran species, with at least 3 loci identified

in species representing the family Hylidae, 2 loci in a species representing Centrolenidae and 3 loci in other species within the family Ranidae, within which *R. pipiens* and *R. temporaria* lie [16].

The MHC class I gene has yet been characterised in a bufonid species. Here we present the characterisation of the cane toad MHC class I gene from a spleen cDNA library and its comparison to other species. Primers were then designed to characterise allelic variation at the α1 domain and α2 domain in the genomic DNA of 25 individuals.

Materials and Methods

Purification of nucleic acids

Xenopus laevis genomic DNA was purified from denuded oocytes using DNeasy Blood and Tissue Kit (QIAGEN) with proteinase K digestion at 56°C for 1.5 hr. Oocyte manipulations followed a protocol approved under the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes. Cane toad toe samples were collected at four locations (Cairns, 16°55′S, 145°46′E, colonised 1936; Normanton, 17°40′S, 141°04′E, colonised 1966; Borroloola, 16°04′S, 136°18′E, colonised 1988; Timber Creek, 15°38′S, 130°28′E, colonised 2006) across the cane toad distribution. Genomic DNA was purified from these toe sample using DNeasy Blood and Tissue Kit (QIAGEN) with proteinase K digestion at 56°C for at least 3 hrs. All protocols were carried out under permits issued by the Northern Territory Parks and Wildlife Commission and with approval of the University of Sydney Animal Care and Ethics Committee (L04/5-2008/2/4790 and L04/11-2006/3/4478). Euthanasia was carried out with pentobarbital with all efforts made to minimise animal suffering.

Total RNA was purified from cane toad tissues using RNeasy Plus mini Kit (QIAGEN) following the manufacturer's protocol. These were DNase treated prior to cDNA synthesis using SuperScript III First-Strand Synthesis System (INVITROGEN).

Construction and screening of cDNA library

A cDNA phage expression library was constructed from a cane toad spleen RNA using the PCR-based SMART cDNA Library Construction Kit (CLONTECH). Size fractionated (>100 bp) cDNA were ligated into the λTriplEx2 arms (CLONTECH) and packaging with a GigaPack III Gold extract kit (STRATAGENE). Primary titres for the unamplified libraries were determined to be approximately 5×10^6 pfu/ml. The libraries were then amplified by transducing *Escherichia coli* XL-1 Blue with phage suspensions according to the manufacturers protocol (CLONTECH).

PCR primers (XelaMHCIa F and XelaMHCIa R; also see table 1) were designed from *X. laevis* MHC Class I sequences available on Genbank to amplify the conserved alpha 3 (exon 4) domain. The 181 bp fragment was amplified from *X. laevis* DNA using PCR in a volume of 20 µl containing 1×PCR buffer (INVITROGEN), 2.5 mM MgCl, 0.2 mM dNTPs, 0.8 mM each primer, and 0.125 U *Taq* polymerase (INVITROGEN).

PCR products were confirmed by electrophoresis on an agarose gel consisting of 2% agarose, 1×TBE (89 mM Tris base, 89 mM boric acid, 2 mM EDTA, pH 8) and 1×SYBRsafe (INVITROGEN) run in 1×TBE electrophoresis buffer. PCR products were excised from the gel and purified using a QIAEX II Gel Purification Kit (QIAGEN) according to the manufacturer's protocol. These PCR products were cloned using the pGEM-T Easy Vector System (PROMEGA) and purified using QIAprep Spin Miniprep Kit (QIAGEN) prior to sequencing on a ABI 3730 sequencer at the Australian Genome Research Facility (AGRF, Westmead, Aus-

tralia). Plasmids confirmed to contain *X. laevis* MHC Class I alpha 3 were used as the DNA template for amplification of the probe for screening the cDNA library.

Plaques were lifted from culture plates using Amersham Hybond N+ hybridisation membranes (GE HEALTHCARE). Membranes were then soaked in denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 7 min, neutralising solution (1.5 M NaCl, 0.5 M Tris, 1 mM EDTA, pH 7.2) for 3 min, washed in 2xSSC (0.3 M NaCl, 0.03 M Trisodium citrate) for 5 min then fixed at 80°C for 2 hr. The MHC Class I alpha 3 probe was radioactively labelled using Random Primed DNA Labelling Kit (ROCHE APPLIED SCIENCE) and [α-^{32}P]dCTP. Unincorporated dNTPs were removed using Illustra ProbeQuant G-50 Micro Columns (GE HEALTHCARE). After overnight hybridization at 55°C, the membranes were washed and subjected to autoradiography overnight or longer at −80°C. Positive plaques were isolated through primary and secondary screenings.

Plaque phages were converted to plasmid by in vivo excision in *E. coli* BM25.8 following the SMART protocol. 5 µl was used to plate on LB agar with 100 µg/ml ampicillin. Isolated colonies were picked and used to inoculate LB broth for overnight incubation followed by plasmid clean up with QIAprep Spin Miniprep Kit (QIAGEN). Purified plasmid DNA was sequenced by AGRF (Westmead, Australia) with the SMART 3′ or 5′ sequencing primers (CLONTECH).

MHC class I allelic diversity

The cane toad MHC class I full length transcript was aligned to the *Xenopus laevis* MHC class I gene (accession no.: M58019; [17]) to predict exons. Two sets of PCR primers were designed to amplify the α1 and the α2 domain using Oligo 6 (MOLECULAR BIOLOGY INSIGHTS, INC.) to investigate allelic diversity in 25 cane toad individuals. These primer binding sites were located in highly conserved regions of the MHC class I regions as identified by sequence alignments with *Xenopus* spp and *Rana* spp. Amplification was carried out in a total volume of 20 µl containing 1× HotStar HiFi PCR Buffer (QIAGEN), 1 µM each primer, 1 U Hotstar HiFidelity DNA Polymerase (QIAGEN) in a thermal cycler with an initial denaturation and activation of the polymerase at 95°C for 5 min, 35 cycles of 94°C for 30 s, 60°C for 1 min, 72°C for 45 s, and a final extension at 72°C for 10 min.

Amplification was confirmed on a 2% agarose/TBE gel. PCR products were excised from the gel and purified using the QIAEX II gel purification kit (QIAGEN). Purified products were ligated into a pGEM-T Easy vector (PROMEGA) prior to transformation of competent JM109 *Escherichia coli* cells (PROMEGA). Transformants were plated onto LB/agar plates supplemented with 100 µg/ml ampicillin, 0.5 mM IPTG and 80 µg/ml X-Gal to allow for blue/white screening and incubated overnight. Positive clones were picked and cultured in LB broth with 50 µg/ml ampicillin overnight prior to plasmid purification using a QIAprep Spin Miniprep Kit (QIAGEN). Purified plasmids were sequenced by AGRF (Westmead, Australia) with the T7 sequencing primer. Twelve clones were collected per PCR, and allele sequences were accepted if they were observed in more than one individual or isolated from two independent PCR reactions.

Sequence analysis

Sequences were edited in BioEdit 7 [18] and confirmed to contain MHC sequences using tblastx (available online at http://blast.ncbi.nlm.nih.gov/Blast.cgi; [19]). Codons were aligned by ClustalW as implemented in MEGA5 [20]. MEGA was also used to construct maximum likelihood phylogenies from the nucleotide sequences, using General Time Reversible model of nucleotide

Table 1. PCR primers.

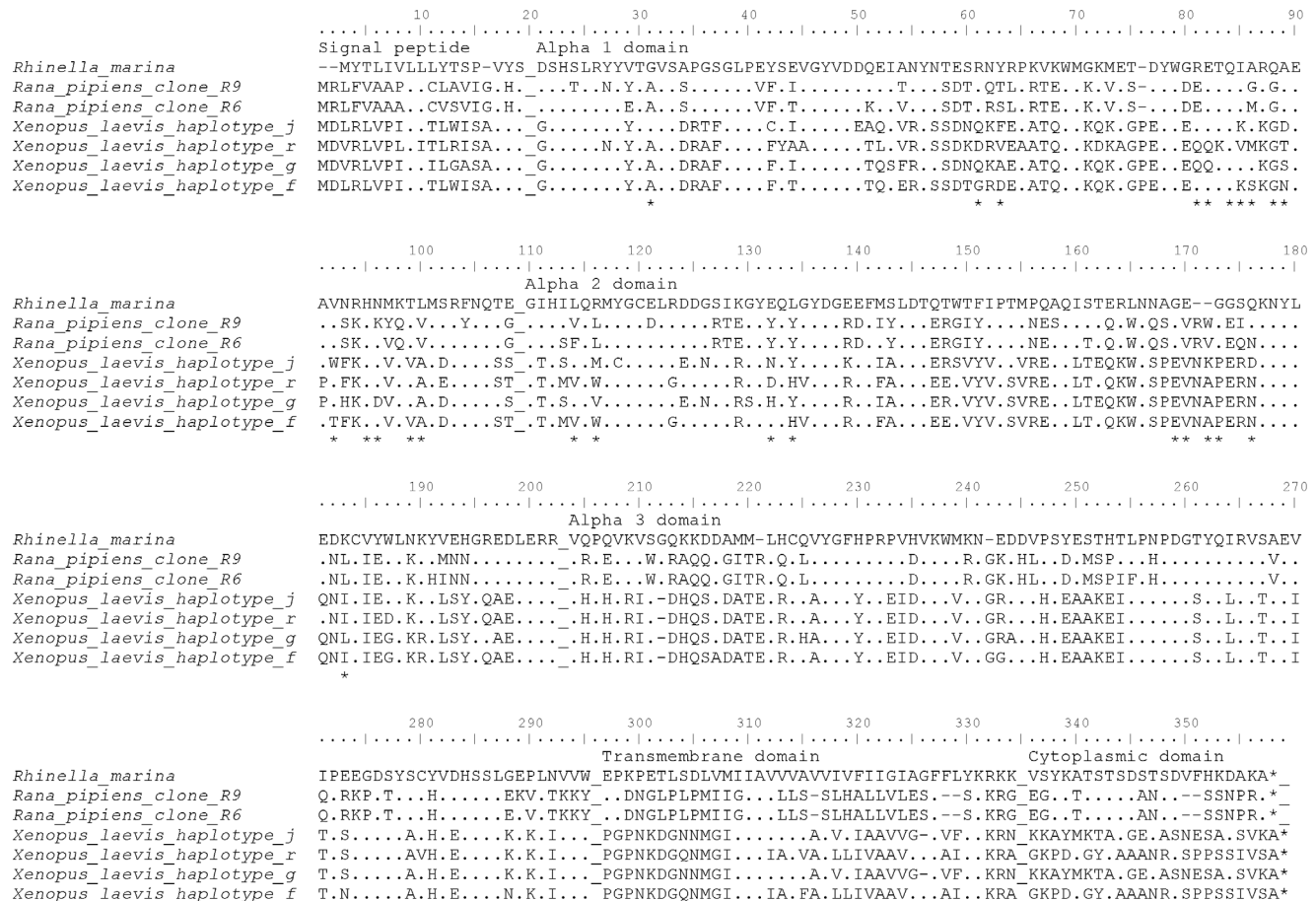| Species	Primer names	Primer sequence (5′-3′)	T_A	Product length
Xenopus laevis	XelaMHCIa F	AATCAGATGACGCCACAGA	56	181
Xenopus laevis	XelaMHCIa R	CCTCATTGGGTGTAATCTCAG	56	
Rhinella marina	Class I_Alpha1domF	ACAGTCACTCTCTGCGTTATT	60	251
Rhinella marina	Class I_Alpha1domR	GTTGAAGCGGCTCATC	60	
				241
Rhinella marina	RhmaA1altF3	TCTGCGTTATTATGTAACTGG	55	
Rhinella marina	Class I_Alpha2domF	GGATGTACGGCTGTGAGCTG	60	243
Rhinella marina	Class I_Alpha2domR	GATCTTCTCTCCCGTGCTCCA	60	

doi:10.1371/journal.pone.0102824.t001

substitution as recommended by Modelgenerator [21] and rooted on their midpoints.

Results and Discussion

Isolation and characterisation of MHC class I alpha chain

Approximately 4.8×10^9 pfus of the cane toad spleen cDNA library were screened with the X. laevis MHC class I α3 domain probe. After primary screening, 18 positive plaques were picked

and cultured for secondary screening. 9 positive plaques were picked, converted to plasmid and sequenced. Partial sequences obtained from five plasmids yielded identical sequences with significant similarity to amphibian MHC class I using tblastx. Two positive plasmids were chosen for complete sequencing of both forward and reverse strands; both plasmids contained identical sequences.

```
                  10        20        30        40        50        60        70        80        90
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
            Signal peptide      Alpha 1 domain
Rhinella_marina              --MYTLIVLLLYTSP-VYS_DSHSLRYYVTGVSAPGSGLPEYSEVGYVDDQEIANYNTESRNYRPKVKWMGKMET-DYWGRETQIARQAE
Rana_pipiens_clone_R9        MRLFVAAP..CLAVIG.H._...T..N.Y.A..S......VF.I.........T...SDT.QTL.RTE..K.V.S-...DE....G.G..
Rana_pipiens_clone_R6        MRLFVAAA..CVSVIG.H._........E.A.S......VF.T......K..V..SDT.RSL.RTE..K.V.S-...DE....M.G..
Xenopus_laevis_haplotype_j   MDLRLVPI..TLWISA..._G......Y....DRTF....C.I.....EAQ.VR.SSDNQKFE.ATQ..KQK.GPE..E....K.KGD.
Xenopus_laevis_haplotype_r   MDVRLVPL.ITLRISA..._G.....N.Y.A..DRAF....FYAA.....TL.VR.SSDKDRVEAATQ..KDKAGPE..EQQK.VMKGT.
Xenopus_laevis_haplotype_g   MDVRLVPI..ILGASA..._G.......Y.A..DRAF....F.I.....TQSFR..SDNQKAE.ATQ..KQK.GPE..EQQ....KGS.
Xenopus_laevis_haplotype_f   MDLRLVPI..TLWISA..._G.......Y.A..DRAF....F.T.....TQ.ER.SSDTGRDE.ATQ..KQK.GPE..E....KSKGN.
                                   *                      *  *         ** *** **

                  100       110       120       130       140       150       160       170       180
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
                          Alpha 2 domain
Rhinella_marina              AVNRHNMKTLMSRFNQTE_GIHILQRMYGCELRDDGSIKGYEQLGYDGEEFMSLDTQTWTFIPTMPQAQISTERLNNAGE--GGSQKNYL
Rana_pipiens_clone_R9        ..SK.KYQ.V...Y...G_....V.L....D.....RTE..Y.Y....RD.IY....ERGIY....NES....Q.W.QS.VRW.EI.....
Rana_pipiens_clone_R6        ..SK..VQ.V...G....SF.L.......RTE..Y.Y....RD..Y..ERGIY....NE...T.Q.W.QS.VRV.EQN....
Xenopus_laevis_haplotype_j   .WFK..V.VA.D....SS_.T.S..M.C.....E.N..R..N.Y....K..IA...ERSVYV..VRE..LTEQKW.SPEVNKPERD....
Xenopus_laevis_haplotype_r   P.FK..V..A.E....ST_.T.MV.W.......G.....R..D.HV...R..FA....EE.VYV.SVRE..LT.QKW.SPEVNAPERN....
Xenopus_laevis_haplotype_g   P.HK.DV..A.D.....S_.T.S..V.......E.N..RS.H.Y....R..IA...EE.VYV..SVRE..LTEQKW.SPEVNAPERN....
Xenopus_laevis_haplotype_f   .TFK..V.VA.D....ST_.T.MV.W.......G.....R....HV...R..FA....EE.VYV.SVRE..LT.QKW.SPEVNAPERN....
                             *   **   **              *  *            * *                    **  **   *

                  190       200       210       220       230       240       250       260       270
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
                          Alpha 3 domain
Rhinella_marina              EDKCVYWLNKYVEHGREDLERR_VQPQVKVSGQKKDDAMM-LHCQVYGFHPRPVHVKWMKN-EDDVPSYESTHTLPNPDGTYQIRVSAEV
Rana_pipiens_clone_R9        .NL.IE..K..MNN........_.R.E...W.RAQQ.GITR.Q.L........D....R.GK.HL..D.MSP...H..........V..
Rana_pipiens_clone_R6        .NL.IE..K.HINN........_.R.E...W.RAQQ.GITR.Q.L........D....R.GK.HL..D.MSPIF.H..........V..
Xenopus_laevis_haplotype_j   QNI.IE..K..LSY.QAE....._.H.H.RI.-DHQS.DATE.R..A...Y..EID....V..GR...H.EAAKEI......S..L..T..I
Xenopus_laevis_haplotype_r   .NI.IED.K..LSY.QAE....._.H.H.RI.-DHQS.DATE.R..A...Y..EID....V..GR...H.EAAKEI......S..L..T..I
Xenopus_laevis_haplotype_g   QNL.IEG.KR.LSY..AE....._.H.H.RI.-DHQS.DATE.R.HA...Y..EID....V..GRA..H.EAAKEI......S..L..T..I
Xenopus_laevis_haplotype_f   QNI.IEG.KR.LSY.QAE....._.H.H.RI.-DHQSADATE.R..A...Y..EID....V..GG...H.EAAKEI......S..L..T..I
                              *

                  280       290       300       310       320       330       340       350
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
                             Transmembrane domain                       Cytoplasmic domain
Rhinella_marina              IPEEGDSYSCYVDHSSLGEPLNVVW_EPKPETLSDLVMIIAVVVAVVIVFIIGIAGFFLYKRKK_VSYKATSTSDSTSDVFHKDAKA*_
Rana_pipiens_clone_R9        Q.RKP.T...H......EKV.TKKY_..DNGLPLPMIIG...LLS-SLHALLVLES.--S.KRG_EG..T.....AN..--SSNPR.*_
Rana_pipiens_clone_R6        Q.RKP.T...H......E.V.TKKY_..DNGLPLPMIIG...LLS-SLHALLVLES.--S.KRG_EG..T.....AN..--SSNPR.*_
Xenopus_laevis_haplotype_j   T.S.....A.H.E....K.K.I..._PGPNKDGNNMGI.......A.V.IAAVVG-.VF..KRN_KKAYMKTA.GE.ASNESA.SVKA*
Xenopus_laevis_haplotype_r   T.S.....AVH.E....K.K.I..._PGPNKDGQNMGI...IA.VA.LLIVAAV...AI..KRA_GKPD.GY.AAANR.SPPSSIVSA*
Xenopus_laevis_haplotype_g   T.S.....A.H.E....K.K.I..._PGPNKDGNNMGI.......A.V.IAAVVG-.VF..KRN_KKAYMKTA.GE.ASNESA.SVKA*
Xenopus_laevis_haplotype_f   T.N.....A.H.E....N.K.I..._PGPNKDGQNMGI...IA.FA.LLIVAAV...AI..KRA_GKPD.GY.AAANR.SPPSSIVSA*
```

Figure 1. Amino acid alignment of anuran MHC class I. ClustalW alignment of MHC class I from Rhinella marina, Rana pipiens (MHC class I clone R6 Genbank accession no.: AF185587 and R9: AF185588) and Xenopus laevis (MHC class I haplotype g: AF188579; haplotype f: AF188580; haplotype r: AF188582 and haplotype j: AF188586); * indicate putative peptide binding residues as predicted from Flajnik et al. 1999.
doi:10.1371/journal.pone.0102824.g001
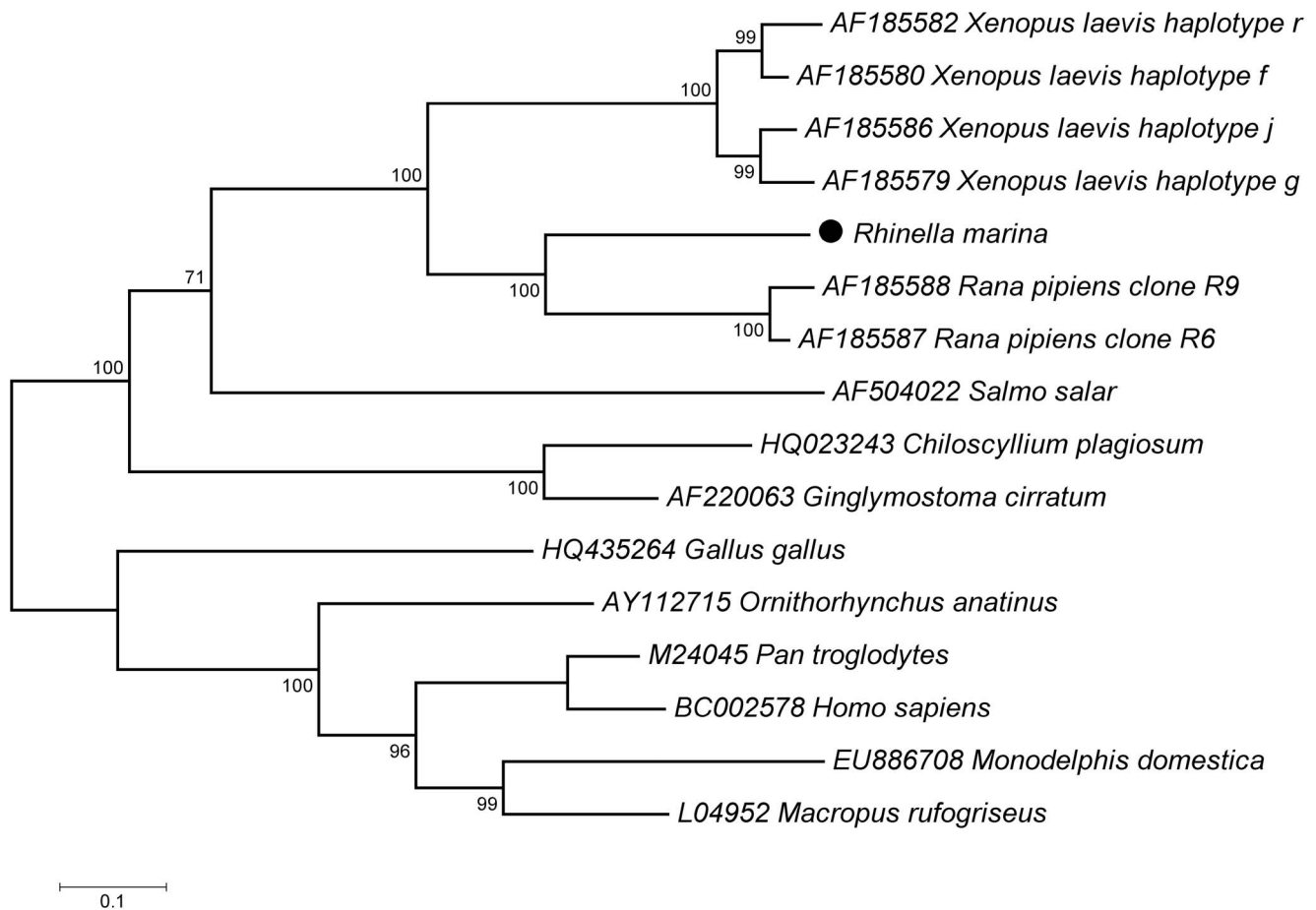
Figure 2. Phylogenetic tree comparing _Rhinella marina_ MHC class I to other vertebrates. Maximum likelihood phylogenetic tree using General Time Reversible model of nucleotide substitution with invariant sites (rooted on midpoint) of MHC class I in cane toad, _Rhinella marina_, African clawed frog, _Xenopus laevis_, northern leopard frog, _Rana pipiens_, Atlantic salmon, _Salmo salar_, whitespotted bamboo shark, _Chiloscyllium plagiosum_, nurse shark, _Ginglymostoma cirratum_, chicken, _Gallus gallus_, platypus, _Ornithorhynchus anatinus_, gray short-tailed opossum, _Monodelphis domestica_, red-necked wallaby, _Macropus rufogriseus_, common chimpanzee, _Pan troglodytes_, and human, _Homo sapiens_. Sequence titles include Genbank accessions.
doi:10.1371/journal.pone.0102824.g002

The cane toad MHC transcript (Genbank accession: KC295548) was aligned to MHC class I genes of other anuran species (_X. laevis_ MHC class I haplotype g: AF188579; haplotype f: AF188580; haplotype r: AF188582 and haplotype j: AF188586; _R. pipiens_ MHC class I clone R6: AF185587 and R9: AF185588). This alignment was used to predict the domains of the expressed MHC class I gene and the putative peptide binding residues (Figure 1).

The sequenced cane toad MHC class I transcript was 1035 bp in length, containing an open reading frame of 344 amino acids, spanning the signal peptide, alpha 1, 2 and 3 domains, and the transmembrane and cytoplasmic domains (Figure 1). It had on average 64% nucleotide similarity and 53% amino acid similarity with the MHC class I of _R. pipiens_, and only 57% nucleotide and 44% amino acid similarity with _X. laevis_. Phylogenetic analysis revealed the cane toad MHC class I clustered together with class I sequences from _R. pipiens_ as a sister clade to MHC class I sequences from _X. laevis_ (Figure 2).

MHC class I α1 and α2 domain polymorphism

Two sets of primers were designed from the full-length transcript to assess the allelic variation in the alpha 1 domain
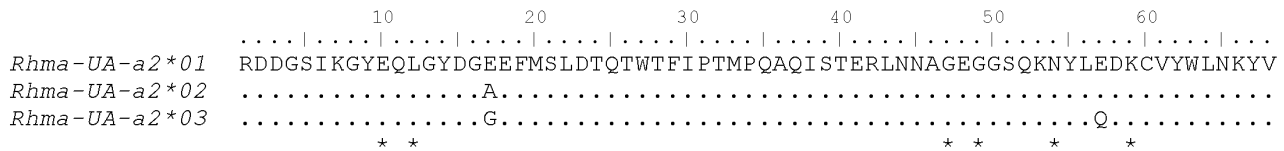
and the alpha 2 domain. Specific residues encoded by these domains are involved with peptide binding and have been shown to be highly polymorphic in many species because they are under the influence of positive selection [22,23]. PCR primers were designed to maximise the number of codons predicted to have peptide binding function from an alignment with _X. laevis_ [14] and primer binding sites were localised to regions that appeared highly conserved across anuran lineages.

The alpha 1 PCR primers (Class I_Alpha1domF and Class I_Alpha1domR; Table 1) amplified a 214 bp fragment that included 14 putative peptide-binding sites. Up to 8 different variants were sequenced in one single individual, indicating that these primers amplify at least 4 loci. From the 25 individuals, 17 alpha 1 variants were characterised (Figure 3; Genbank accessions: KC295549-KC295565). These variants share between 73% to 99.5% nucleotide sequence similarity and between 54.9% and 98.5% amino acid sequence similarity. Allele variant _Rhma-UA-a1*01_ was identical to the full length transcript and had between 94.4–98.1% sequence similarity to _Rhma-UA-a1*02_ and _*03_. These variants form a strongly supported phylogenetic clade (Figure 4) and a maximum of two alleles were found in each individual, supporting our conclusion that these alleles are from the _Rhma-UA_ locus characterised from the cDNA library. The

```
Alpha 1 domain

                  10        20        30        40        50        60        70
             ....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
Rhma-UA-a1*01 YVTGVSAPGSGLPEYSEVGYVDDQEIANYNTESRNYRPKVKWMGKMETDYWGRETQIARQAEAVNRHNMKT
Rhma-UA-a1*02 ...............................................E......................
Rhma-UA-a1*03 ......................G...V.............EN.G.E.......................
Rhma-I-a1*01  .K.A.........S.........V...S..GR........E.V.PG..KEQ....KGT...Y...VR.
Rhma-I-a1*02  .KMA...............R..T...S..KILKG.....E.VDPG..EHN...TKSN...F...VR.
Rhma-I-a1*03  .T.A.........F.I.....R..V...S...RMK......E.V.PE..E....KQKP....F..DVR.
Rhma-I-a1*04  .Y.A.........F.I.....RK..Y..S..GRMK.N....E.L.PG..EGQ..R.KSQ.PLY...VR.
Rhma-I-a1*05  .Y.A.........F.I.....RK..Y..S..GRMK.N....E.L.PG...GQ..R.KSQ.PLY...VR.
Rhma-I-a1*06  .Y.A.........F.I.....RK..Y..S..GRMK.N....E.L.LG..EGQ..S.KSQ.PLY...VR.
Rhma-I-a1*07  .Y.A.........F.I.....RK..Y..S...RMK.NM...E.L.PG..EGQ..S.KSQ.PLY...VR.
Rhma-I-a1*08  .Y.A.........F.S.........V...S..GR.......E.VDPE..E......KGN.V.F.R.LR.
Rhma-I-a1*09  .Y.A.........S.....R..V...S..GR.......E.VDPG..KGQ....KGN...Q...VR.
Rhma-I-a1*10  .Y.A....E....S.........V...S..GR.......E.V.PE..E....KGN...F...VR.
Rhma-I-a1*11  .Y.A....E....S.........V...S..GR.......E.V.PE..E....KGN...F.R.LR.
Rhma-I-a1*12  .Y.........A......R..V...S.KGKEQ......E.V.PG..EEQ....KGN...F...VR.
Rhma-I-a1*13  .Y.........A......R..V...S.KGKEQ......E.V.PE..EEQ....KGN...F...VR.
Rhma-I-a1*14  .Y.A.........F.I.....R..V...S..GRMK.N....E.V.PE..EGQ..R.KGN...F...VR.
                *                                     *  *        ** *** **  *  **  *
```

```
Alpha 2 domain

                  10        20        30        40        50        60
             ....|....|....|....|....|....|....|....|....|....|....|....|....|...
Rhma-UA-a2*01 RDDGSIKGYEQLGYDGEEFMSLDTQTWTFIPTMPQAQISTERLNNAGEGGSQKNYLEDKCVYWLNKYV
Rhma-UA-a2*02 ...............A..................................................
Rhma-UA-a2*03 ...............G................................Q............
                 *  *                                    *  *    *    *
```

Figure 3. Amino acid alignment of MHC class I alpha 1 domain and alpha 2 domain variants. Putative peptide binding sites are indicated with * under each alignment.

other alpha 1 allele variants may represent other alleles from other classical loci, alleles from non-classical loci, or pseudogenes. As classification and locus assignment of these alleles was beyond the scope of this study, we named them *Rhma-I-a1* ("I" for MHC class I, "a1" for alpha 1 domain, and numbered in order of characterisation).

The alpha 2 PCR primers (Class I_Alpha2domF and Class I_Alpha2domR; Table 1) amplified a 202 bp fragment, containing 6 putative peptide-binding sites. A maximum of 2 variants were isolated per individual, indicating the primers were specific for one class I locus. A total of 3 alleles were sequenced from 25 individuals (Figure 3; Genbank accession: KC295569-KC295571), which had between 98% and 99% nucleotide sequence similarity, and between 97% and 98.5% amino acid sequence similarity. Allele variant *Rhma-UA-a2*01* was identical to the full length transcript and other allele variants were numbered in order of characterisation.

Phylogenetic trees constructed for the MHC class I alpha 1 and alpha 2 domains (Figure 4) show the cane toad allele variants forming a monophyletic clade. In both cases, they form a sister clade to class I sequences isolated from the Hylid frog, *Smilisca phaeota*. Both species represent the Hyloidea superfamily that diversified during the Paleogene, and are estimated to have last shared a common ancestor 66 MYA [24].

Classification of *Rhma-UA-a1* allele variants

Our results indicate a large expansion of class I loci in the cane toad. Our aims were to characterise a classical class I locus for future population genetic surveys, as these typically are the most polymorphic class I loci. As such, we focused on the full length

transcript characterised from the cDNA library and the alleles *Rhma-UA-a1*01*, *02*, and *03* that belong to this locus.

As a preliminary investigation into whether *Rhma-UA-a1*01-03* were indeed classical class I alleles, we compared the alpha 1 variants expressed in 7 different tissues, including liver, lung, kidney, heart, ventral skin, dorsal skin and spleen from one individual to the alpha 1 variants amplified from its genomic DNA. We expected that the classical class I alpha 1 variants would be highly expressed in all tissues, and as a result we would observed these variants at the highest frequency. Twenty clones were sequenced per tissue and 40 clones were sequenced from the genomic DNA. Results are summarised in table 2.

Eleven alpha 1 variants in total were sequenced across all tissues and genomic DNA. An additional 3 alpha 1 variants (*Rhma-I-a1*15–17*; Genbank accession: KC295566-KC295568), previously uncharacterised, were isolated in this sequencing. The *Rhma-UA-a1*01* variant was the most highly represented sequence isolated across all tissues. The remaining alpha 1 variants had variable expression patterns throughout the tissues, indicating that these variants may be non-classical. Although the *Rhma-I-a1*02*, *09*, *12* variants showed a wide tissue distribution, their sequence frequency were lower than those of *Rhma-UA-a1*01*, supporting this notion. Further work to classify the non-classical allele variants and assign loci will provide valuable insight into the organisation of the cane toad MHC class I. With 10 putative non-classical alpha 1 variants sequenced from this one individual, it appears as though there are at least 5 class I loci in the cane toad in addition to the classical class I *Rhma-UA* locus. Characterisation of full-length transcripts will help to ascertain whether all these putative non-
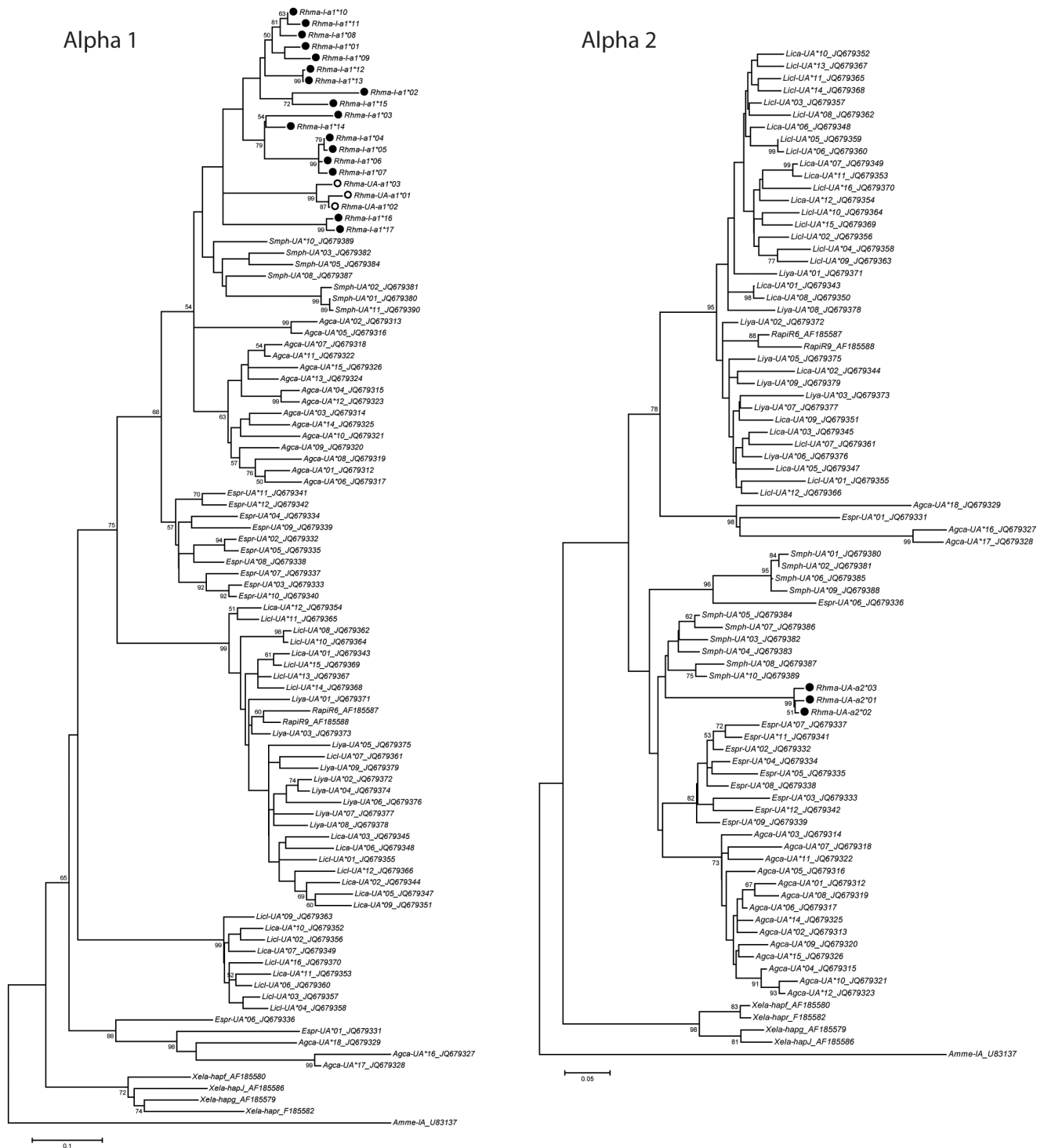
Figure 4. Phylogenetic trees for MHC class I alpha 1 domain and alpha 2 domain. Maximum likelihood phylogenetic trees using General Time Reversible model of nucleotide substitution with invariant sites (rooted on midpoint) comparing *Rhinella marina* (Rhma) to other anuran species (*Xenopus laevis*, Xela; *Silurana tropicalis*, Sitr; *Rana temporaria*, Rate; *Agalychnis callidryas*, Agca; *Espadarana prosoblepon*, Espr; *Smilisca phaeota*, Smph; *Lithobates catesbeianus*, Lica; *L. clamitans*, Licl; *L. yavapaiensis*, Liya) and axolotl (*Ambystoma mexicanium*, Amme). Sequence titles include Genbank accessions.
doi:10.1371/journal.pone.0102824.g004

classical allele variants are functional genes, or pseudogenes, and allow locus assignment. Expression analysis will also help to classify these alleles as either classical or non-classical, based on their differential expression patterns.

It is worth noting that our primers could preferentially amplify the classical class I locus over other variants. The primers were designed from the full-length MHC class I transcript, which shared over 90% similarity to any *Rhma-UA-a1* allele. If there were

Table 2. Numbers of allele variants sequenced from cDNA from various tissues and the genomic DNA of one cane toad individual from Normanton, Australia.

	Rhma UA-a1*01	Rhma-I-a1 *02	Rhma-I-a1 *03	Rhma-I-a1 *04	Rhma-I-a1 *07	Rhma-I-a1 *08	Rhma-I-a1 *09	Rhma-I-a1 *12	Rhma-I-a1 *15	Rhma-I-a1 *16	Rhma-I-a1 *17
gDNA	14	4	1	3	2	5	1	2		3	3
Spleen	7	1	1		1	3	4			3	
Ventral Skin	14	1				1	2	1	1	1	
Lung	12	1	3				4				
Liver	8	2	2			2	2	3			
Dorsal Skin	17	1		1				1			
Kidney	7		3			2	4	1	1		
Heart	18	1						1			

Forty clones were sequenced from genomic DNA and 20 clones were sequenced from each tissue type. Recombinant clones were identified and removed from genomic DNA (2), kidney (2), liver (1) and spleen (3) records.

doi:10.1371/journal.pone.0102824.t002

mismatches in the primer-binding site, amplification of the *Rhma-I-a1* variants could occur at a lower frequency and lead to an underrepresentation of the putative non-classical allele variants in our cloning and sequencing.

We have subsequently sequenced a cane toad spleen transcriptome (Illumina HiSeq paired-end library, ABySS assembly, custom blast database, blastn and tblastx queries with our full length transcript and our allele variants). Contigs with significant similarity to the cane toad alpha 1 allele variants were extracted and aligned. This alignment revealed that the sequence at the ClassI_Alpha1domR priming site was 100% conserved between all UA allele variants. The sequence at the ClassI_Alpha1domF priming site was also highly conserved with a few exceptions. One C-T mismatch was identified on allele variants *Rhma-I-a1*03*, *12, and *08*, however as these appeared in our cloning and sequencing, we can assume that our primers do bind and amplify these allele variants. We also observed a previously uncharacterised allele variant in the transcriptome with two C-T mismatches at the priming site, which may imply lower primer binding and subsequently low or no amplification of this allele variant. No other class I allele variants were observed in the transcriptome using these methods. Comparison of the alpha 2 domain primer-binding sites reveals that the primer sequences were highly specific to the *Rhma-UA* locus, which serves to explain why these primers did not amplify other MHC class I loci and the absence of non-classical alpha 2 domain allele variants in our cloning and sequencing results.

A new forward primer specific for the *Rhma-UA* alpha 1 locus was designed (RhmaAIaltF3; Table 1) that allows the amplification of a 204 bp fragment (together with the original reverse primer, Class I_Alpha1domR).

Genetic variation of MHC class I α1 and α2 domains

The MHC class I is typically highly polymorphic in natural populations, but we only found 3 alpha 1 variants and 3 alpha 2 variants at *Rhma-UA* in 25 Australian cane toads. This is exceptionally low MHC class I diversity when compared to similar studies in other species (Table 3). Furthermore, these sequences were highly similar with between 94.4% and 98.1% nucleotide similarity at the alpha 1 domain and between 98% and 99.5% at the alpha 2 domain.

This severe depletion of MHC class I variation is the likely result of the history of serial translocations events, which gave rise to the modern Australian cane toad population. The repeated founder events drastically reduced genetic diversity across the entire cane toad genome. Only a single mitochondrial haplotype is found in the Australian population, which was identical to that found in their source population of Hawaii [25].

Relaxed balancing selection on the MHC may have further depleted variation in the MHC of the Australian cane toad. Prior to their introduction, bufonid toads were absent from Australia, and presumably so were the pathogens that have coevolved with them. The most effective pathogens of the Australian cane toad may be those that were co-introduced, but surveys of Australian and Hawaiian cane toad have demonstrated that they lack many of the parasites common in the natural populations in South America [26].

One South American parasite that was introduced to Australia with the cane toad is a lungworm, *Rhabdias pseudosphaerocephala*, which causes morbidity and mortality in a wide range of toad life-history stages [27,28]. This parasite lags behind the cane toad invasion front [29], and thus the selection pressure it mediates across the population would be variable. Indeed it appears as though selection for dispersal traits has led to reduced immune

Table 3. Comparison of allelic diversity at the MHC class I (N is number of individuals studies and NA is number of alleles).

Species and Citation	Class I domain/s studied	# loci	N	NA
Cane toad (*Rhinella marina*) this study	**Alpha 1**	**1**	**25**	**3**
	Alpha 2	**1**	**25**	**3**
Common frog (*Rana temporaria*) [15]	Alpha 1 and 2	1	66	129
Sockeye salmon (*Oncorhynchus nerka*) [32]	Alpha 2	1	>100	34
Atlantic salmon (*Salmo salar*) [33]	Alpha 1,2, 3	1	82	11
Banded houndshark (*Triakis scyllium*) [34]	Alpha 1,2, 3 (Trsc-UAA)	1	22	29
	Alpha 1,2, 3 (Trsc-UBA)	1	22	6
Blue tit (*Cyanistes caeruleus*) [35]	Alpha 2	4	20	17
House Sparrow (*Passer domesticus*) [36]	Alpha 2	≥3	8	20
House Sparrow (*Passer domesticus*) [37]	Alpha 2	≥3	10	27

doi:10.1371/journal.pone.0102824.t003

function on the invasion front [30,31]. Further research will investigate whether this translates into reduced selection pressure for MHC polymorphism leading to depletion of MHC diversity over the course of the cane toad invasion.

In conclusion, we characterised a single classical MHC class I gene in the cane toad and we infer an expansion of non-classical MHC class I loci from sequencing of alpha 1 domain allele variants. We found low MHC variability in Australian cane toads; the likely result of repeated bottleneck events during its history of introductions. This low MHC variability may indicate an immunological weakness of this invasive species, which may give hope to the development of a bio-control strategy.

Author Contributions

Conceived and designed the experiments: ML KB. Performed the experiments: ML. Analyzed the data: ML. Contributed reagents/materials/analysis tools: RS. Contributed to the writing of the manuscript: ML RS KB.

References

1. Easteal S (1981) The history of introductions of *Bufo-marinus* (Amphibia, Anura) - A natural experiment in evolution. Biological Journal of the Linnean Society 16: 93–&.
2. Frost DR, Grant T, Faivovich J, Bain RH, Haas A, et al. (2006) The amphibian tree of life. Bulletin of the American Museum of Natural History: 8–370.
3. Urban MC, Phillips BL, Skelly DK, Shine R (2007) The cane toad's (*Chaunus Bufo marinus*) increasing ability to invade Australia is revealed by a dynamically updated range model. Proceedings of the Royal Society B-Biological Sciences 274: 1413–1419.
4. Klein J (1986) Natural History of the Major Histocompatibility Complex. New York: Wiley & Sons.
5. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, et al. (1987) The foreign antigen-binding site and T-cell recognition regions of class-I histocompatibility antigens. Nature 329: 512–518.
6. Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. Annual Review of Genetics. Palo Alto: Annual Reviews.pp. 121–152.
7. Braud VM, Allan DSJ, McMichael AJ (1999) Functions of nonclassical MHC and non-MHC-encoded class I molecules. Current Opinion in Immunology 11: 100–108.
8. Bahram S (2000) MIC genes: from genetics to biology. Advances in immunology 76: 1–60.
9. Ohta Y, Goetz W, Hossain MZ, Nonaka M, Flajnik MF (2006) Ancestral organization of the MHC revealed in the amphibian *Xenopus*. Journal of Immunology 176: 3674–3685.
10. Shum BP, Avila D, Dupasquier L, Kasahara M, Flajnik MF (1993) Isolation of a classical MHC class-I cDNA from an amphibian - Evidence for only one class-I locus in the *Xenopus* MHC. Journal of Immunology 151: 5376–5386.
11. Flajnik MF, Kasahara M, Shum BP, Saltercid L, Taylor E, et al. (1993) A Novel Type of Class-I Gene Organization in Vertebrates - A Large Family of Non-MHC-Linked Class-I Genes is Expressed at the RNA Level in the Amphibian *Xenopus* Embo Journal 12: 4385–4396.
12. Goyos A, Sowa J, Ohta Y, Robert J (2011) Remarkable Conservation of Distinct Nonclassical MHC Class I Lineages in Divergent Amphibian Species. Journal of Immunology 186: 372–381.

13. Salter-Cid L, Nonaka M, Flajnik MF (1998) Expression of MHC class Ia and class Ib during ontogeny: High expression in epithelia and coregulation of class Ia and lmp7 genes. Journal of Immunology 160: 2853–2861.
14. Flajnik MF, Ohta Y, Greenberg AS, Salter-Cid L, Carrizosa A, et al. (1999) Two ancient allelic lineages at the single classical class I locus in the *Xenopus* MHC. Journal of Immunology 163: 3826–3833.
15. Teacher AGF, Garner TWJ, Nichols RA (2009) Evidence for Directional Selection at a Novel Major Histocompatibility Class I Marker in Wild Common Frogs (*Rana temporaria*) Exposed to a Viral Pathogen (Ranavirus). Plos One 4.
16. Kiemnec-Tyburczy KM, Richmond JQ, Savage AE, Lips KR, Zamudio KR (2012) Genetic diversity of MHC class I loci in six non-model frogs is shaped by positive selection and gene duplication. Heredity 109: 146–155.
17. Flajnik MF, Canel C, Kramer J, Kasahara M (1991) Evolution of the Major Histocompatibility Complex - Molecular-cloning of the Major Histocompatibility Complex Class-I from the Amphibian *Xenopus*. Proceedings of the National Academy of Sciences of the United States of America 88: 537–541.
18. Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.
19. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Searsh Tool. Journal of Molecular Biology 215: 403–410.
20. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28: 2731–2739.
21. Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McInerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. Bmc Evolutionary Biology 6.
22. Edwards SV, Hedrick PW (1998) Evolution and ecology of MHC molecules: from genomics to sexual selection. Trends in Ecology & Evolution 13: 305–311.
23. Piertney SB, Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. Heredity 96: 7–21.
24. Roelants K, Gower DJ, Wilkinson M, Loader SP, Biju SD, et al. (2007) Global patterns of diversification in the history of modern amphibians. Proceedings of

the National Academy of Sciences of the United States of America 104: 887–892.

25. Slade RW, Moritz C (1998) Phylogeography of *Bufo marinus* from its natural and introduced ranges. Proceedings of the Royal Society of London Series B-Biological Sciences 265: 769–777.

26. Barton DP (1997) Introduced animals and their parasites: The cane toad, *Bufo marinus*, in Australia. Australian Journal of Ecology 22: 316–324.

27. Dubey S, Shine R (2008) Origin of the parasites of an invading species, the Australian cane toad (*Bufo marinus*): are the lungworms Australian or American? Molecular Ecology 17: 4418–4424.

28. Kelehear C, Webb JK, Shine R (2009) *Rhabdias pseudosphaerocephala* infection in Bufo marinus: lung nematodes reduce viability of metamorph cane toads. Parasitology 136: 919–927.

29. Phillips BL, Kelehear C, Pizzatto L, Brown GP, Barton D, et al. (2010) Parasites and pathogens lag behind their host during periods of host range advance. Ecology 91: 872–881.

30. Llewellyn D, Thompson MB, Brown GP, Phillips BL, Shine R (2012) Reduced investment in immune function in invasion-front populations of the cane toad (*Rhinella marina*) in Australia. Biological Invasions 14: 999–1008.

31. Brown GP, Shilton C, Phillips BL, Shine R (2007) Invasion, stress, and spinal arthritis in cane toads. Proceedings of the National Academy of Sciences of the United States of America 104: 17698–17700.

32. McClelland EK, Ming TJ, Tabata A, Miller KM (2011) Sequence analysis of MHC class I alpha2 from sockeye salmon (*Oncorhynchus nerka*). Fish & shellfish immunology 31: 507–510.

33. Grimholt U, Drablos F, Jorgensen SM, Hoyheim B, Stet RJM (2002) The major histocompatibility class I locus in Atlantic salmon (*Salmo salar* L.): polymorphism, linkage analysis and protein modelling. Immunogenetics 54: 570–581.

34. Okamura K, Ototake M, Nakanishi T, Kurosawa Y, Hashimoto K (1997) The most primitive vertebrates with jaws possess highly polymorphic MHC class I genes comparable to those of humans. Immunity 7: 777–790.

35. Schut E, Aguilar JD, Merino S, Magrath MJL, Komdeur J, et al. (2011) Characterization of MHC-I in the blue tit (*Cyanistes caeruleus*) reveals low levels of genetic diversity and trans-population evolution across European populations. Immunogenetics 63: 531–542.

36. Bonneaud C, Sorci G, Morin V, Westerdahl H, Zoorob R, et al. (2004) Diversity of Mhc class I and IIB genes in house sparrows (*Passer domesticus*). Immunogenetics 55: 855–865.

37. Borg AA, Pedersen SA, Jensen H, Westerdahl H (2011) Variation in MHC genotypes in two populations of house sparrow (*Passer domesticus*) with different population histories. Ecology and evolution 1: 145–159.